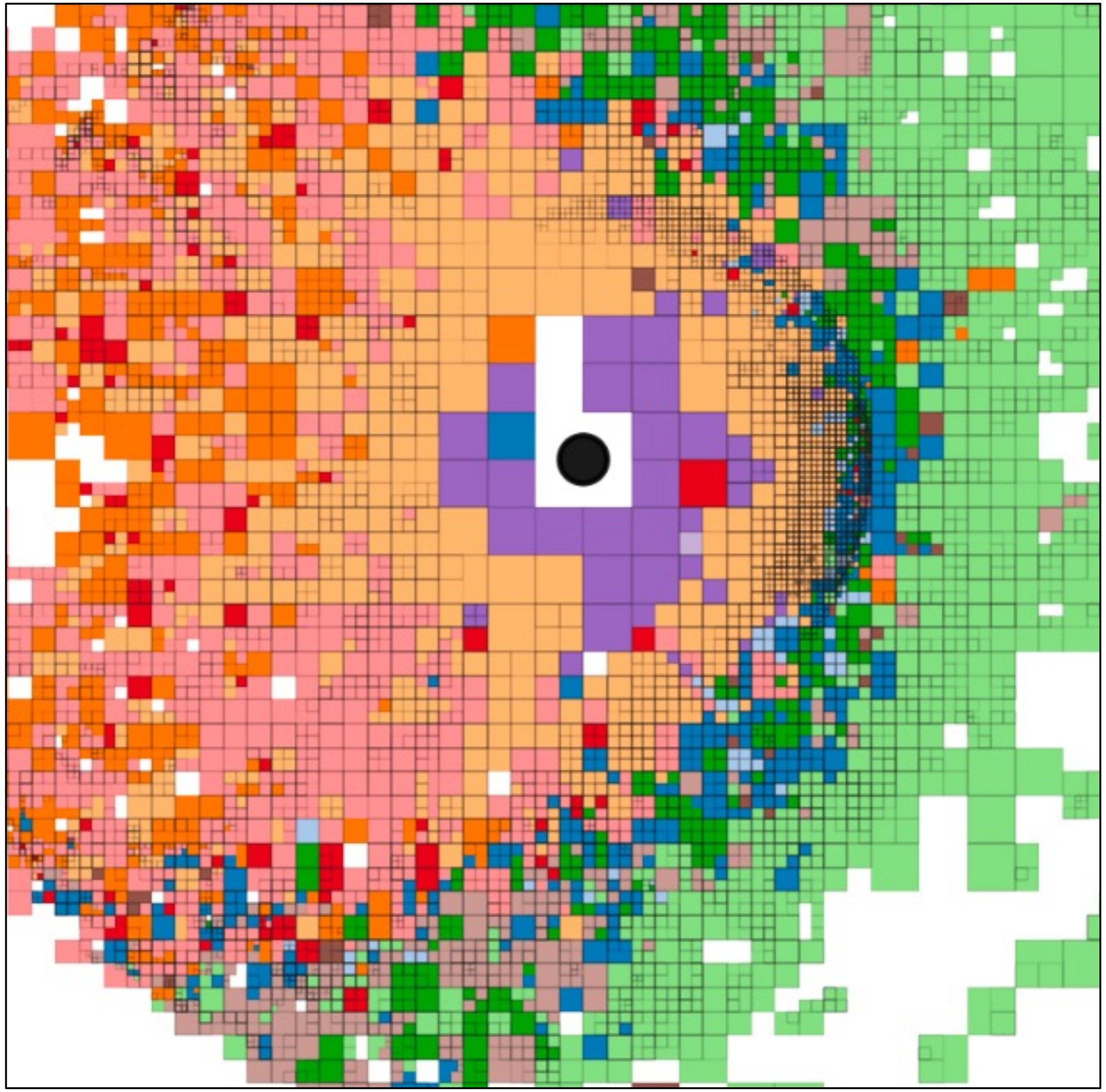
A Generative Atlas of Earth's Magnetosphere for Space Weather

From MMS observations to probabilistic mapping of plasma regimes

Stefano Markidis, Jonah Ekelund, Luca Pennati

KTH Royal Institute of Technology



ASAP
Automatics in Space Exploration

Can we build a probabilistic atlas of Earth's magnetosphere for space weather?

- We now have enough in-situ measurements to test this idea seriously.
- MMS provides a strong first proof of concept for a data-driven magnetosphere atlas.
- The opportunity is to move from a single-mission research prototype to a multi-mission ESA capability.

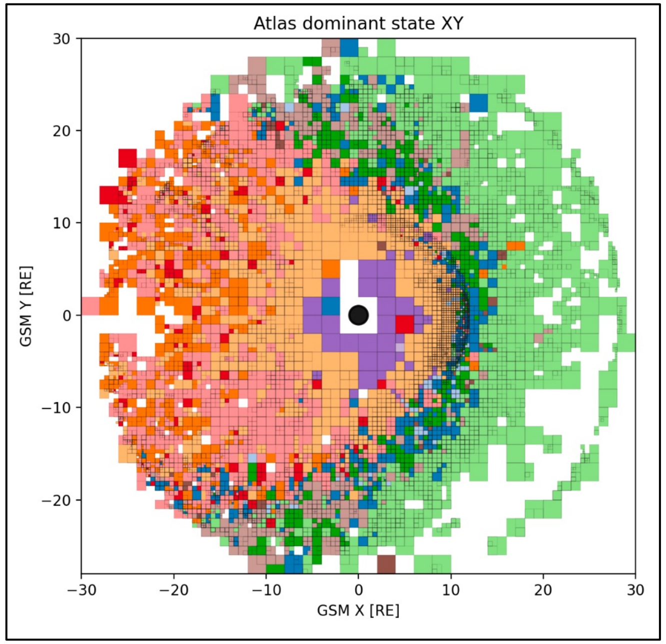
The question: what can current data already tell us?

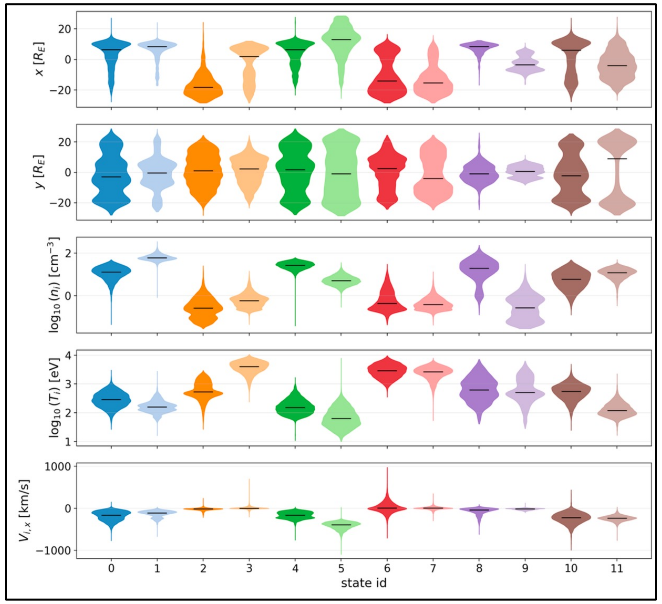
- How far can current in-situ data take us toward mapping magnetospheric plasma regimes?
- Can we condition that map on space-weather activity, starting with Dst?
- The goal is not only region labeling, but estimating probability distributions of plasma states.
- This would support:
  - regime identification,
  - anomaly detection,
  - data assimilation,
  - probabilistic space-weather modelling.

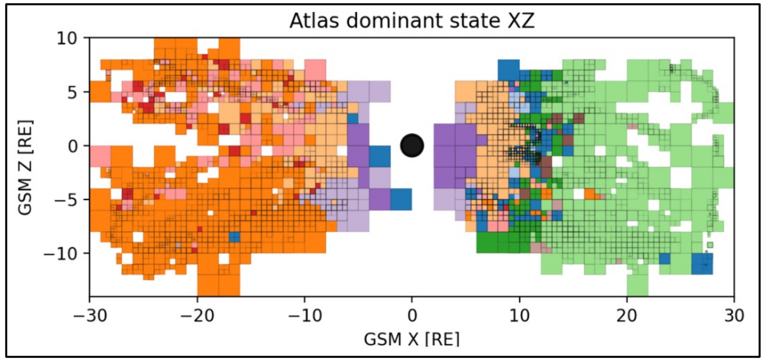
MMS already gives a proof of concept

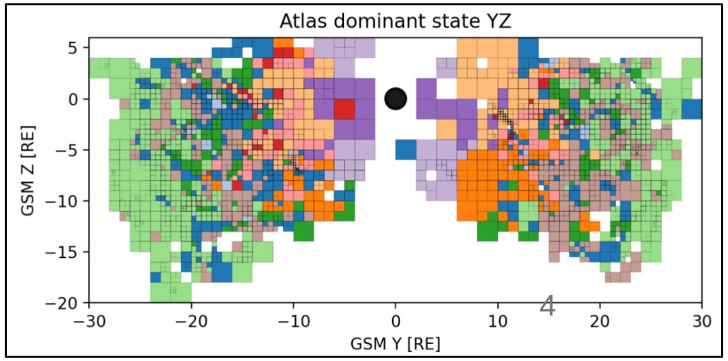
- We built the prototype from about 660,000 MMS1 state windows over 10 years.
- Each state summarizes local plasma conditions over a 2-minute window.
- The method learns 12 discrete plasma regimes without expert labels.
- Python pipeline ingesting MMS1 dataset, processing, and applying the VQE

Item	Value
Total states	660,340
Time span (UTC)	2015-09-20 to 2025-10-30
Unique days	2,614
State windowing	$\Delta t = 2\,\mathrm{min}$, cadence $10\,\mathrm{s}$, stride $1\,\mathrm{min}$ (50% overlap)
Spatial extent (R_E)	$x \in [-28.37, 28.48]$, $y \in [-28.46, 28.70]$, $z \in [-20.38, 10.38]$
Quiet-time fraction	86.2% with $\lvert Dst \rvert \leq 20\,\mathrm{nT}$
Features ($d = 14$)	$\bar{x}$ and $\sigma(x)$ for $x \in \{\lvert B \rvert, B_z, n_i, T_i, \lvert V_i \rvert, V_{i,x}, \beta_i^*\}$
Octree support	Adaptive 3D octree on median window position; base cell size $2\,R_E$, refine if > 25 states/cell; 65,352 leaf cells



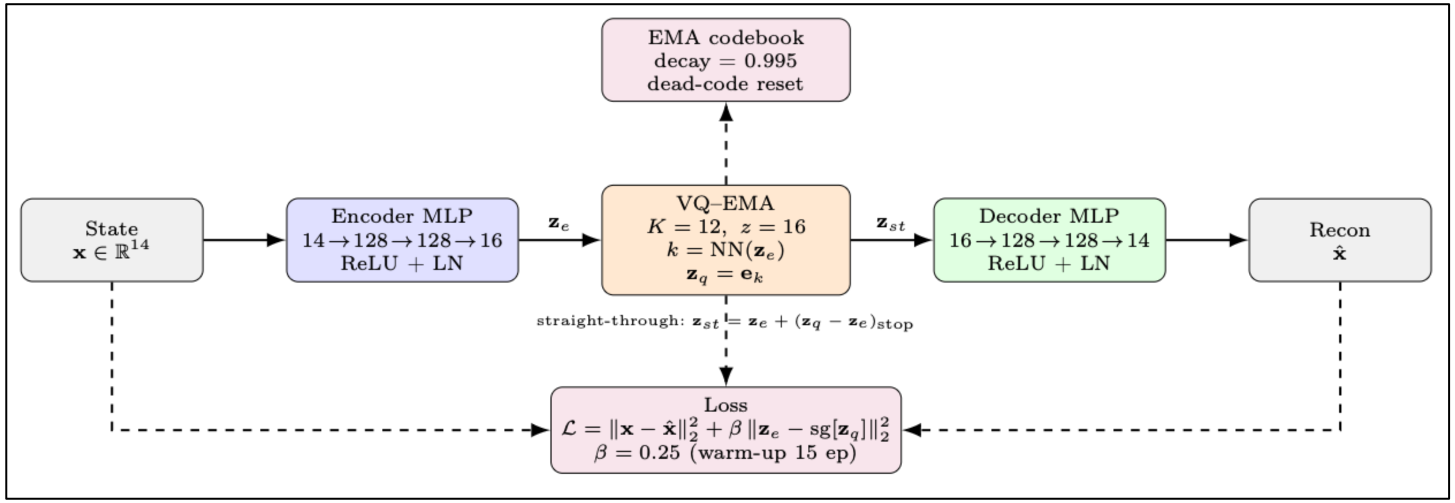






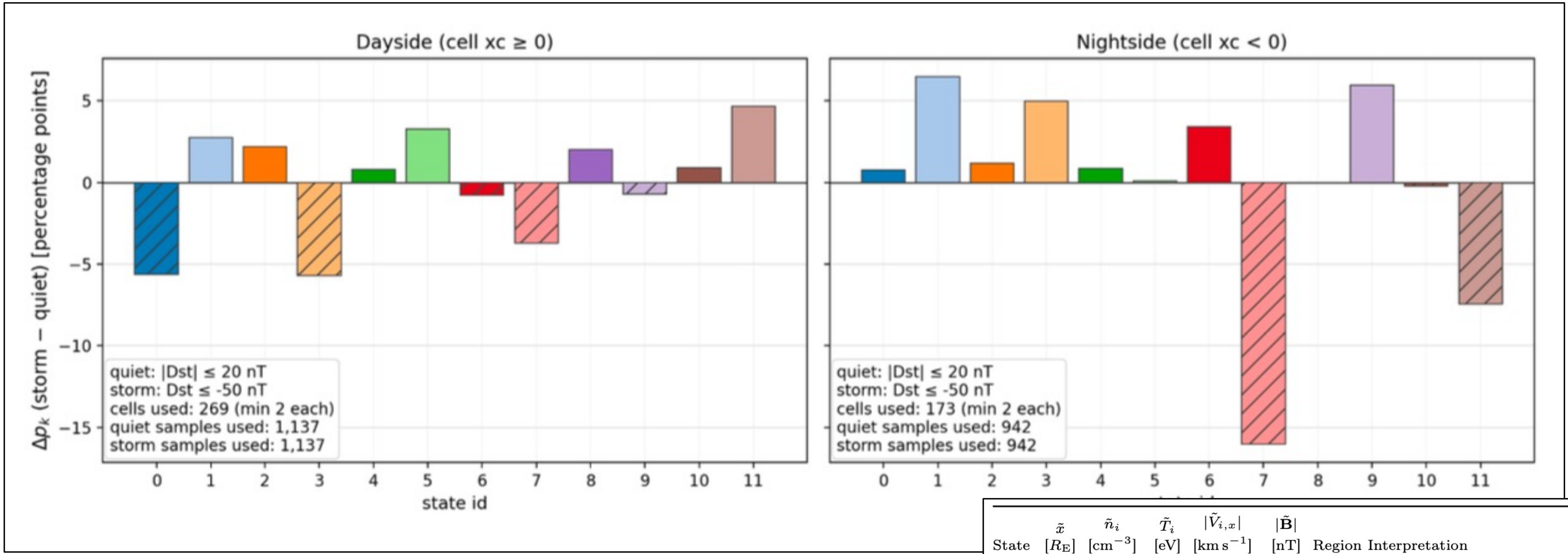
What makes it possible

- Spacecraft data are strongly non-uniform in space because they follow orbital tracks.
- We address this with an adaptive 3D octree:
    - high resolution where data are dense,
    - coarse resolution where data are sparse.
- We then use a VQ-VAE generative model to learn discrete plasma regimes.
- Each regime is represented not by one number, but by a distribution of plasma conditions.
- This makes the atlas both probabilistic and interpretable.



State	$\tilde{x}$ [R_{E}]	$\tilde{n}_i$ [cm^{-3}]	$\tilde{T}_i$ [eV]	$\lvert\tilde{V}_{i,x}\rvert$ [$\mathrm{km\,s^{-1}}$]	$\lvert\tilde{\mathbf{B}}\rvert$ [nT]	Region Interpretation
0	6.4	13.0	286	163	19.7	MSH (hot/processed; boundary-adjacent)
1	8.3	59.9	160	109	24.5	Dense MSH (shocked solar wind)
2	-18.2	0.3	528	15	20.3	PSBL / lobe-adjacent tail transition
3	1.9	0.6	4056	5	34.9	Hot plasma sheet / inner hot magnetosphere
4	6.3	27.2	154	159	19.3	MSH
5	13.1	5.1	63	392	5.6	Solar wind
6	-14.0	0.4	2896	8	20.5	Plasma sheet (dynamic/bursty)
7	-15.3	0.4	2637	5	13.4	Tail plasma sheet
8	8.5	19.6	623	41	41.5	Compressed/stagnation sheath near dayside boundary (mixing)
9	-3.4	0.3	510	14	100.2	Strong-field, very low-β environment (lobe-like / near-Earth dipolar high-latitude)
10	6.0	6.0	561	220	15.9	Magnetopause boundary layer / hot sheath
11	-3.9	12.2	119	235	14.2	Flank MSH (downstream)

Use case: Dst-conditioned magnetosphere response



- We compare regime occupancy during quiet and storm-time conditions using Dst.
- The atlas captures meaningful space-weather behavior:
    - reduced tail plasma sheet occupancy on the nightside,
    - increased strong-field / low-β states,
    - more energetic plasma populations near Earth.

State	$\tilde{x}$ [R_E]	$\tilde{n}_i$ [cm^{-3}]	$\tilde{T}_i$ [eV]	$\lvert\tilde{V}_{i,x}\rvert$ [km s^{-1}]	$\lvert\tilde{\mathbf{B}}\rvert$ [nT]	Region Interpretation
0	6.4	13.0	286	163	19.7	MSH (hot/processed; boundary-adjacent)
1	8.3	59.9	160	109	24.5	Dense MSH (shocked solar wind)
2	-18.2	0.3	528	15	20.3	PSBL / lobe-adjacent tail transition
3	1.9	0.6	4056	5	34.9	Hot plasma sheet / inner hot magnetosphere
4	6.3	27.2	154	159	19.3	MSH
5	13.1	5.1	63	392	5.6	Solar wind
6	-14.0	0.4	2896	8	20.5	Plasma sheet (dynamic/bursty)
7	-15.3	0.4	2637	5	13.4	Tail plasma sheet
8	8.5	19.6	623	41	41.5	Compressed/stagnation sheath near dayside boundary (mixing)
9	-3.4	0.3	510	14	100.2	Strong-field, very low-β environment (lobe-like / near-Earth dipolar high-latitude)
10	6.0	6.0	561	220	15.9	Magnetopause boundary layer / hot sheath
11	-3.9	12.2	119	235	14.2	Flank MSH (downstream)

From research prototype to ESA capability

Infrastructure

- Build a harmonized unified multi-mission database with magnetosphere "states"
  - MMS, Cluster, THEMIS + others
    - Larger 3D coverage
    - More coverage in storm periods
- Standardized:
  - Coordinates
    - GSM vs. GSE
  - plasma & field parameters
    - Units
  - metadata and uncertainty

Integration

- Connect existing ESA assets and datasets
- Define common plasma regimes and probabilities
- Enable interoperability by providing "most likely plasma states" conditioned on *X* for:
  - modelling
  - simulation
  - data assimilation

Operational value

- Provide a real-time probabilistic context of the magnetosphere given a number of conditional variables
- Support:
  - anomaly detection
  - space-weather forecasting
- Deliver benchmark products for the community

A new way to represent the magnetosphere

- The magnetosphere can be described as a set of probabilistic plasma regimes, learned directly from data
- This provides a data-driven, observation-based view complementary to physics models
- It enables a shift from deterministic regions → to probabilistic, condition-dependent states
- We can now move from describing the magnetosphere to statistically characterizing its behavior.

A Generative Atlas of the Earth's Magnetosphere

Stefano Markidis, Jonah Ekelund, and Luca Pennati

KTH Royal Institute of Technology, Stockholm, Sweden

Abstract. We propose a methodology to identify and characterize different plasma environments observed by magnetospheric mission spacecraft instruments, creating a data-driven magnetosphere atlas. To enable an accurate analysis under highly non-uniform sampling, we design a 3D adaptive octree data structure that refines magnetospheric regions with high sampling counts. A variational autoencoder that leverages the octree structure allows us to automatically identify different regions of the Earth's magnetosphere. The model also generates per-region feature distributions, improving interpretability by summarizing both typical conditions and variability. We apply this methodology to 10 years of plasma observations from NASA's Magnetospheric Multiscale (MMS) mission, identify known magnetospheric regions, and create an atlas of plasma environments in space. The resulting atlas can be used to better understand magnetospheric plasma regions, detect anomalies, automatically generate labels for space applications, and study the magnetospheric response across geomagnetic activities.

Keywords: Earth magnetosphere · atlas · unsupervised learning · VQ-VAE · generative models · adaptive grids

To appear in ICCS 2026

Acknowledgements

This work is supported by the European Union's HORIZON RIA via ASAP - Automatics in Space under the grant agreement no. 101082633.

We acknowledge the MMS mission team and the instrument teams for producing the MMS data products used in this study. MMS data are publicly available via NASA's MMS Science Data Center.

ASAP
Automatics in Space Exploration